

# Sequence of *Escherichia coli* O128 antigen biosynthesis cluster and functional identification of an $\alpha$ -1,2-fucosyltransferase

Jun Shao, Mei Li, Qiang Jia, Yuquan Lu<sup>1</sup>, Peng George Wang<sup>\*,2</sup>

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

Received 20 June 2003; revised 8 August 2003; accepted 11 August 2003

First published online 9 September 2003

Edited by Giovanni Cesareni

**Abstract** O128 is one of the most common atypical enteropathogenic *Escherichia coli* isolated from diarrhea patients worldwide. The primary structure of *E. coli* O128 repeat units has previously been determined as  $\rightarrow 3$ - $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)- $\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 6)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$  pentasaccharide. Here we report the complete sequencing of *E. coli* O128 antigen biosynthesis gene cluster and its flanking regions. Comparative sequence analysis revealed the expected O128 antigen process genes, GDP-fucose biosynthesis genes and four potential glycosyltransferase genes responsible for the assembly of *E. coli* O128 antigen repeats. *WbsJ* was shown to encode an  $\alpha$ -1,2-fucosyltransferase by enzymatic assays and nuclear magnetic resonance spectroscopy analysis.

© 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

**Key words:** Lipopolysaccharide;  $\alpha$  (1 $\rightarrow$ 2) fucosyltransferase; *Escherichia coli* O128 gene cluster

## 1. Introduction

Enteropathogenic *Escherichia coli* (EPEC) is a leading cause of infantile diarrhea in developing countries. In industrialized countries, the frequency of these organisms has decreased, but they continue to be an important cause of diarrhea [1]. EPEC are divided into typical and atypical subgroups based on the presence of the adherence factor (EAF) plasmid [2]. The most studied EPEC strains belong to a series of O-antigenic serotypes known as EPEC O serogroups. As one of the most common atypical EPEC, *E. coli* O128 strains have been isolated from diarrhea patients worldwide, for example, in Brazil [3], England [4], Peru [5], Italy [6] and Singapore [7].

Bacterial lipopolysaccharides (LPS) typically consist of a hydrophobic domain known as lipid A, a core oligosaccha-

ride, and a distal polysaccharide (O antigen) that contributes major antigenic variability to the cell surface [8]. The O polysaccharide repeat units are synthesized from sugar nucleotides by glycosyltransferases that are often soluble or peripheral membrane proteins, and assembled on the membrane-bound carrier, undecaprenyl phosphate (undP), at the inner face of the cytoplasmic membrane [8]. The genes involved in the biosynthesis of O antigens in *E. coli* are generally clustered adjacent to the *gnd* gene between the colanic acid and *his* operons [9].

The need for potent vaccines against infectious diseases is becoming more urgent as antibiotic-resistant pathogens continue to emerge. Vaccines consisting of carbohydrates coupled to a protein carrier have been proven effective for the prevention of invasive bacterial disease [10]. However, these glycoconjugates prepared by using purified polysaccharides or degraded oligosaccharides are sometimes contaminated with other bacterial components. They may also have ill-defined structures due to multiple coupling sites [11]. The molecular size of carbohydrate antigens is another major concern in designing glycoconjugate vaccines [12,13]. Large polysaccharides may decrease the solubility of glycoconjugates, and size-unconstrained antigens may yield heterogeneous conjugates with the potential for immunogenic inconsistency [14]. Therefore, the structurally defined glycoconjugates generated from synthetic carbohydrate antigens with desired molecular sizes are especially attractive in our battle against infectious diseases. The exploration of these well-structured glycoconjugate vaccines has been hindered largely by the technical difficulties in chemical preparation of oligosaccharides.

The primary structure of LPS O-specific polysaccharide (Fig. 1) from *E. coli* O128 has been established [15]. The immunodominant part of *E. coli* O128 polysaccharide was determined as  $\beta$ -D-GalNAc-(1 $\rightarrow$ 6)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal trisaccharide by ELISA-inhibition study [16]. It can be envisioned that the efficient synthesis of the oligosaccharide would pave the way for the development of a well-structured glycoconjugate vaccine against infantile diarrhea *E. coli* O128. Compared to chemical synthesis, the enzymatic approach utilizing glycosyltransferases has been proven more efficient in the production of complex carbohydrates. Considering the extremely narrow substrate specificity of most known glycosyltransferases, molecular characterization of the biosynthetic pathway of O128 antigen repeat units is undoubtedly a crucial step towards large-scale synthesis of the immunodominant oligosaccharide. In this study, we sequenced the *E. coli* O128 O-antigen gene cluster and its flanking regions to elucidate the genetic basis of O128 antigen biosynthesis. Sequence analysis

\*Corresponding author. Fax: (1)-313-577 9241.

E-mail address: [pwang@chem.wayne.edu](mailto:pwang@chem.wayne.edu) (P.G. Wang).

<sup>1</sup> Present address: Department of Social and Environmental Medicine F1, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan.

<sup>2</sup> Present address: The Ohio State University, Department of Biochemistry and Chemistry, 876 Biological Sciences Building, 484 West 12th Avenue, Columbus, OH 43210, USA. Fax: (1)-614-292 6773. Email: [pgwang@chemistry.ohio-state.edu](mailto:pgwang@chemistry.ohio-state.edu).

**Abbreviations:** EPEC, enteropathogenic *Escherichia coli*; LPS, lipopolysaccharide; undP, undecaprenyl phosphate; GalNAc, *N*-acetylgalactosamine; Gal, galactose; Fuc, fucose; PCR, polymerase chain reaction; GST, glutathione *S*-transferase

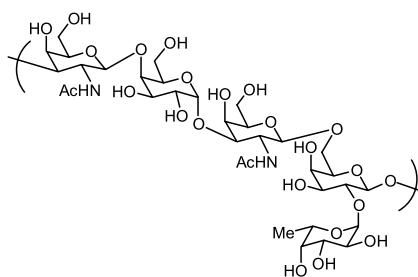


Fig. 1. Structure of *E. coli* O128 antigen.

revealed five genes involved in GDP-fucose (GDP-Fuc) biosynthesis and all three O128 antigen processing genes. More important, we assigned four potential glycosyltransferases responsible for the assembly of *E. coli* O128 antigen repeats. The  $\alpha$ -1,2-fucosyltransferase function of gene *WbsJ* was unambiguously determined by enzymatic assays and nuclear magnetic resonance (NMR) spectroscopy analysis.

## 2. Materials and methods

### 2.1. Bacterial strains and plasmids

*E. coli* O128:B12:H (ATCC 12810) was obtained from American Type Culture Collection (Rockville, MD, USA). *E. coli* TOP10 strain and linearized pCR4Blunt-TOPO vector were from Invitrogen (Carlsbad, CA, USA). The pGEX-4T-1 plasmid was from Amersham Biosciences (Piscataway, NJ, USA). GST-Bind kit was from Novagen (Madison, WI, USA).

### 2.2. Polymerase chain reaction (PCR) and random shotgun cloning

The chromosomal DNA of *E. coli* O128 was prepared by ELU-QUIK DNA purification kit from Schleicher and Schuell Bioscience (Keene, NH, USA). Long-range PCR was carried out by using MasterAmp<sup>®</sup> extra-long PCR kit from Epicentre (Madison, WI, USA). A pair of primers: F (5'-GCTTCTTGCCGATATTTGGC) and R (5'-CCGACGGACTGATCCGGTGATT) were designed based on the highly conserved sequences of *E. coli* *galF* and *HisI* genes, respectively. A total of 15 PCR products were pooled to minimize the effect of PCR errors. After gel purification, the PCR products were sheared by nebulizers (Invitrogen) and cloned into pCR4Blunt-TOPO vector to generate a bank (280 colonies) according to the instruction manual of the TOPO shotgun cloning kit (Invitrogen).

### 2.3. Sequencing and analysis

The plasmids for sequencing were prepared with QIAprep 96 Turbo Miniprep kit by using the BioRobot 9600 of Qiagen (Valencia, CA, USA). Sequencing was performed by an Applied Biosystems 3700 DNA analyzer (Foster City, CA, USA) at Wayne State University DNA sequencing facility. The whole DNA sequence was assembled by the Phred/Phrap package from the Genome Center of the University of Washington. The open reading frames (ORFs) were recognized and selected by using the ORF Finder from the National Center for Biotechnology Information (NCBI). The BLAST database was used to search for sequence homologies and the amino acid sequence alignments were performed using CLUSTAL W [17]. Potential transmembrane segments were identified by TMpred [18].

### 2.4. Cloning and expression of *wbsJ* gene

The *WbsJ* gene was cloned into the *Bam*HI and *Xho*I sites of pGEX-4T-1 plasmid with an N-terminal glutathione *S*-transferase (GST) tag. A pair of primers: P<sub>1</sub> 5'-CGCGGATCCATGGAAGT-TAAAATTATTGGGG-GGCT/P<sub>2</sub> 5'-CGGAATTCCTCGAGTCAT-AATTTTACCCACGATTTCG, with *Bam*HI and *Xho*I restriction sites (underlined) incorporated into P<sub>1</sub> and P<sub>2</sub>, respectively, were used for PCR amplification of *wbsJ* gene from chromosomal DNA. The plasmid pGEX/*WbsJ* was transformed into *E. coli* BL21 (DE3) for expression. The 1-l Luria–Bertani broth culture was induced with 1 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (30°C, 10 h). The pelleted cells were stored at –80°C until needed.

The cell pellet was resuspended in GST Bind/Wash buffer (4.3 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.47 mM KH<sub>2</sub>PO<sub>4</sub>, 137 mM NaCl, 2.7 mM KCl, pH 7.3) and disrupted by sonication (Branson Sonifier 450, VWR Scientific) on ice. The lysate was cleared by centrifugation and loaded onto a GST-Bind column (Novagen). After washing by five volumes of the same buffer, the protein was eluted by GST elution buffer (50 mM Tris–HCl, 10 mM reduced glutathione, pH 8.0).

### 2.5. Fucosyltransferase activity assay

The enzyme assays were performed at 37°C for 2 h in a final volume of 50  $\mu$ l containing 20 mM Tris–HCl (pH 7.0), 10 mM MnCl<sub>2</sub>, 0.3 mM GDP-L-[U-<sup>14</sup>C]Fuc (7000 cpm), 10 mM acceptor galactose (Gal) and various amounts of purified enzyme. Acceptor was omitted for blank. The reaction was terminated by adding 100  $\mu$ l of ice-cold 0.1 M EDTA. Dowex 1 $\times$ 8-200 chloride anion exchange resin was then added. After centrifugation, the radioactivity of the supernatant was counted in a liquid scintillation counter (Beckmann LS-3801 counter). One unit of activity is defined as the amount of enzyme that catalyzes the transfer of 1  $\mu$ mol of Fuc from GDP-Fuc to acceptor per minute at 37°C.

### 2.6. NMR spectroscopy

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded by 500 MHz Varian VXR500 NMR spectrometer. Product structure was identified by one-dimensional (selective COSY, relay COSY, and NOE) and two-dimensional (COSY, HMQC, NOESY, and HMBC) <sup>1</sup>H/<sup>13</sup>C NMR. The oligosaccharide product was repeatedly dissolved in D<sub>2</sub>O and lyophilized before NMR spectra were recorded at 303 K in a 5 mm tube.

### 2.7. Nucleotide sequence accession number

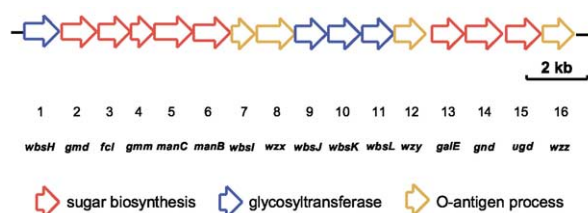
The DNA sequence of *E. coli* O128 antigen biosynthesis gene cluster has been deposited in GenBank under the accession number AY217096.

## 3. Results and discussion

The total of 19013 bp *E. coli* O128 antigen biosynthesis gene cluster, which covers a continuous region from the end of *galF* gene to the middle of *hisI*, was amplified by long-range PCR from the chromosomal DNA of *E. coli* O128 strain and completely sequenced. There are 16 predicted ORFs within the gene cluster (Fig. 2), all in the same transcriptional direction from *galF* to *hisI* gene. The *E. coli* O128 antigen repeat units consist of one L-Fuc, two Gal and two *N*-acetylgalactosamine (GalNAc) residues. It is anticipated that the cluster would contain genes responsible for GDP-L-Fuc synthesis, genes coding glycosyltransferases to assemble the O unit and genes for O-antigen process. The amino acid

Table 1  
Summary of *E. coli* O128 antigen biosynthesis genes

ORF	Name	No. of AA	Putative function
1	<i>wbsH</i>	373	Galactosyltransferase
2	<i>gmd</i>	374	GDP-mannose 4,6-dehydratase
3	<i>fcl</i>	321	GDP-Fuc synthetase
4	<i>gmm</i>	167	GDP-mannose mannosyl hydrolase
5	<i>manC</i>	482	GDP-mannose pyrophosphorylase
6	<i>manB</i>	459	Phosphomannomutase
7	<i>wbsI</i>	185	O-acetyltransferase
8	<i>wzx</i>	480	O-antigen flippase
9	<i>wbsJ</i>	283	Fucosyltransferase
10	<i>wbsK</i>	288	Glycosyltransferase
11	<i>wbsL</i>	299	Glycosyltransferase
12	<i>wzy</i>	347	O-antigen polymerase
13	<i>galE</i>	339	UDP-glucose C4 epimerase
14	<i>gnd</i>	468	Gluconate 6-P dehydrogenase
15	<i>ugd</i>	388	UDP-glucose dehydrogenase
16	<i>wzz</i>	337	O-antigen chain length determinant

Fig. 2. *E. coli* O128 antigen biosynthesis gene cluster.

sequences were used to BLAST search previously characterized genes for possible functions.

It has been well established that GDP-L-Fuc is synthesized from GDP-D-mannose via a three-step pathway catalyzed by two enzymes, GDP-mannose 4,6-dehydratase (*gmd*) and GDP-Fuc synthetase (*fcl*), common to both prokaryotic and eukaryotic cells [19]. Three more enzymes, phosphomannose isomerase (*manA*), phosphomannomutase (*manB*) and GDP-mannose pyrophosphorylase (*manC*), are needed to generate GDP-mannose from fructose 6-phosphate. Comparative sequence analysis revealed five genes involved in GDP-Fuc biosynthesis (Table 1), including a GDP-mannose mannosyl hydrolase (*gmm*), which catalyzes the hydrolysis of GDP- $\alpha$ -D-mannose to yield GDP and  $\beta$ -D-mannose. This enzyme has been suggested to participate in the regulation of cell wall biosynthesis by influencing the concentration of GDP-mannose in the cell [20].

The *E. coli* O polysaccharides are normally synthesized in a Wzy-dependent pathway, which consists of the assembly of O-repeat units on the inner face of cytoplasmic membrane, the exportation by O-antigen flippase Wzx and the polymerization by polymerase Wzy on the periplasmic face of cytoplasmic membrane [8]. The deduced amino acid sequence of *orf8* predicts a 55.3 kDa protein with 12 transmembrane segments. It shows low-level homology with some membrane proteins such as Ta0703 from *Thermoplasma acidophilum* and Wzx from *Streptococcus pneumoniae*. All these proteins have similar functions, the translocation lipid-linked polysaccharide repeat units across the cytoplasmic membrane. Therefore, *orf8* was named *wzx*. The *orf12* encodes a protein with 10 transmembrane segments and two periplasmic regions of 40–50 amino acids. BLAST search indicates that it has low homology with a group of polysaccharide polymerases, including Wzy from *E. coli* O104 and *Shigella boydii* O6. So *orf12* was designated as *wzy*, O-antigen polymerase gene. The *orf16* was named *wzz* because it has 98% identity with chain length determinant proteins from *E. coli* O157 and K12 strains.

Table 2

500 MHz NMR  $^1\text{H}/^{13}\text{C}$  chemical shift values of GalOMe and Fuc $\alpha$ 1,2GalOMe

Residue	Group	H/C $\delta$ (ppm)	
		GalOMe	Fuc $\alpha$ 1,2GalOMe
GalOMe	1	4.15/103.95	4.24/102.85
	2	3.33/70.86	3.34/78.34
	3	3.48/72.90	3.66/72.06
	4	3.75/68.79	3.74/68.91
	5	3.53/75.27	3.51/75.05
	6a	3.63/61.11	3.62/61.09
	6b	3.58/61.11	3.58/61.09
Fuc	OMe	3.41/57.29	3.41/57.26
	1'		4.95/100.22
	2'		3.60/73.38
	3'		3.70/69.70
	4'		3.64/68.57
	5'		4.08/67.01
	6'		1.04/15.34

Five sugar residues are present in *E. coli* O128 antigen repeat unit. The first step in O polysaccharide biosynthesis involves the formation of an und-PP-linked sugar by transferring a sugar 1-phosphate residue to undP [8]. It has been shown that the *wecA* gene, responsible for adding the first sugar to lipid carrier undP and initiating the O unit biosynthesis, resides outside the *E. coli* O-antigen gene clusters [21]. Therefore, we expected to find only four glycosyltransferases within O128 antigen biosynthesis gene cluster. The WecA protein homolog was proposed to transfer the first sugar 1-phosphate to undP on the cytoplasmic face of the cell membrane.

BLAST search revealed that the *orf1*, named *wbsH*, is homologous to a family of bacterial galactosyltransferases. The closest homologs are WbgM (58% identity) from *E. coli* O55 and WbnE (36% identity) from *E. coli* O113 strain. The WbgM has been previously suggested to be a galactosyltransferase that is responsible for the Gal $\alpha$ 1,3GalNAc linkage in *E. coli* O55 repeat units [22]. Tentatively, we assigned WbsH as an  $\alpha$  (1 $\rightarrow$ 3) galactosyltransferase that makes the Gal $\alpha$ (1 $\rightarrow$ 3)GalNAc linkage.

The *orf7* (*wbsI*) predicts a 20.5 kDa protein with high level of homology to a large family of microbial acetyl transferases. O polysaccharide acetylation has been found in a wide range of pathogenic bacteria including *E. coli* [23,24]. It is suggested that the O-acetyl groups play an essential role in the immunogenicity and pathogenesis of these species [25–27]. Thus, WbsI is likely an O-acetyltransferase that is responsible for the acetylation of *E. coli* O128 O polysaccharide, though the precise sugar residues modified are yet to be determined.

The WbsJ protein encoded by *orf9* has a conserved domain

	Motif I	Motif II	Motif III
<i>E. coli</i> WbsJ	GGLGNQMFQYATAFIAIK	CSLHTRRGDYVS	MLMSKCKNNIISNSSFSWAAALN
<i>E. coli</i> WbgN	GGLGNQMFQYAKGYAESV	VSVHTRRGDYVS	MULMSNAKYHIIANSSFSWGAALK
<i>H. pylori</i> Fut2	GGLGNQMFQYAFKSLQK	VFVHTRRGDYVG	MLMQSCKEGIIANSTYSWAAAYLI
<i>B. fragilis</i> WcfB	GRLGNLFFQIATAASLTQ	VSTHVRRGDYLK	LYTQSLCTHNIISNSSFSWGAALN
<i>V. cholerae</i> WblA	GGLGNCLFQYAVGRMAI	VSTHVRRGDYLN	LMLMCCQCHNIVANSSFSWAAALN
<i>L. lactis</i> EpsH	GNLGNCLFIYATAKKIQK	ICVSIIRRGDYVD	VQLMSSCKHFLVLSNSSFSWUTFLS
Human FUT2	GRLGNQMGAYATLYALAK	VGVHVRRGDYVH	FALLTQCNHTIMTIGTFGIWAAAYLT
Bovine FUT2	GRLGNQMGAYATLYALAK	VGVHVRRGDYVH	FALLTQCNHTIMTIGTFGIWAAAYLA
Mouse FUT2	GRLGNQMGAYATLFAALR	VGVHVRRGDYVR	IALLMQCNHTVITLGTFGIWAAAYLT
Rat FUT2	GRLGNQMGAYATLFAALR	VGVHVRRGDYVH	FALLTQCNHTIMTIGTFGIWAAAYLA

Fig. 3. Conserved motifs found in  $\alpha$  (1 $\rightarrow$ 2) fucosyltransferases. The white letters with black background represent identical amino acids, while the black letters with gray background represent similar amino acids conserved in all aligned sequences.



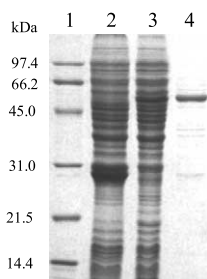


Fig. 4. SDS-PAGE for expression of GST-WbsJ fusion protein. 1, Low-range molecular weight markers (Bio-Rad); 2, cell with pGEX-4T-1 plasmid; 3, cell with pGEX/WbsJ plasmid; 4, purified protein.

from glycosyltransferase family 11. This family mostly consists of  $\alpha$  (1  $\rightarrow$  2) fucosyltransferases from different organisms (<http://afmb.cnrs-mrs.fr/CAZY>). The WbsJ shows a low level of similarity in amino acid sequence to its prokaryotic and eukaryotic counterparts, with 35% identity to *E. coli* WbgN, 30% identity to *Helicobacter pylori* FucT2, 27% identity to *Yersinia enterocolitica* WbcH and 22% identity to human FUT2. However, multiple sequence alignment revealed that WbsJ shares three highly conserved motifs with most of those  $\alpha$  (1  $\rightarrow$  2) fucosyltransferases (Fig. 3). The motif II, which is also shared by the  $\alpha$  (1  $\rightarrow$  6) fucosyltransferases from both prokaryotic and eukaryotic origins, has been suggested to be a part of the GDP-Fuc binding domain [28]. Therefore, we propose that *wbsJ* encodes a fucosyltransferase that transfers a Fuc residue to Gal in  $\alpha$  (1  $\rightarrow$  2) linkage.

The N-terminus of WbsK encoded by *orf10* has a conserved domain from glycosyltransferase family 2. The glycosyltransferases of this family transfer sugar residue with the inversion of configuration. The amino acid sequence of WbsK shares 37%, 34%, 31% and 29% identities with *Lactobacillus delbrueckii* EpsH, *Streptococcus agalactiae* CpsVII, *Haemophilus influenzae* Rd LgtD [29] and *S. pneumoniae* Cps14J, respectively. All these enzymes have been identified as  $\beta$  (1  $\rightarrow$  4) or  $\beta$  (1  $\rightarrow$  3) GalNAc (or Gal) transferases, which participate in the biosynthesis of microbial O-specific or capsular polysaccharides. Thus, we assigned WbsK as the *N*-acetylgalactosaminyltransferase that transfers a GalNAc residue to Gal in  $\beta$  (1  $\rightarrow$  4) linkage. WbsL shows low similarity to several group 1 putative bacterial polysaccharide transferases, and it is likely to be the remaining  $\beta$  (1  $\rightarrow$  6) *N*-acetylgalactosaminyltransferase.

To investigate the function of the *wbsJ* gene, we cloned this

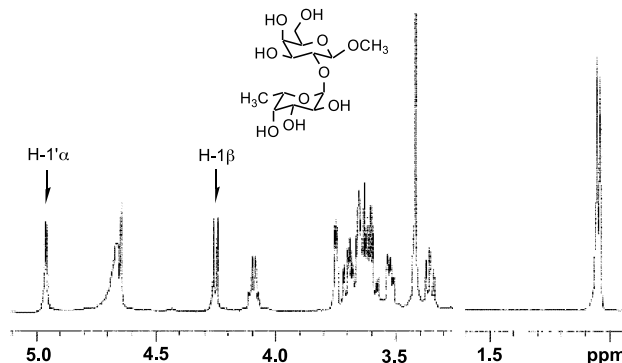


Fig. 5. 500 MHz proton one-dimensional NMR spectrum of Fuc $\alpha$ 1,2GalOMe.

gene into pGEX-4T-1 plasmid. The WbsJ was expressed in *E. coli* BL21 (DE3) strain as a GST fusion protein and purified by one-step GST affinity column. The fusion protein has an apparent molecular weight of 59 kDa estimated by SDS-PAGE (Fig. 4), similar to the theoretical value (59 806.2 Da) calculated from its predicted amino acid sequence. Preliminary kinetic analysis shows that the fusion enzyme has  $K_m$  values of 138  $\mu$ M and 120 mM to the donor GDP-Fuc and acceptor methyl  $\beta$ -Gal, respectively. By using purified enzyme, mg-scale synthesis was performed with GDP-Fuc and methyl  $\beta$ -Gal as donor and acceptor substrates. Total 4.8 mg of disaccharide product was purified by Bio-Gel P2 gel filtration to allow  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectroscopy analysis.

Signals were found for a newly introduced Fuc residue in the NMR spectrum (Fig. 5) of the disaccharide product. A coupling constant of 4 Hz for the H-1' resonance of Fuc showed that this residue was linked in  $\alpha$ -anomeric configuration (Table 2). The significant downfield increments in the chemical shift of the C-2 resonance of Gal ( $\Delta\delta$  = 7.48 ppm), but not in other resonances, indicated that the Fuc residue had been introduced to C-2 of Gal residue. The linkage was further confirmed by the two crosspeaks in the HMBC spectrum (data not shown). One crosspeak was between the C-1' of Fuc and the H-2 of the Gal residue, the other was between the H-1' of Fuc and the C-2 of the Gal residue. Therefore, the structure of the disaccharide product was confirmed to be Fuc $\alpha$ 1,2GalOMe.

In conclusion, the complete *E. coli* O128 antigen biosynthesis gene cluster was sequenced. All the genes required for the biosynthesis of O antigen were identified by comparative sequence analysis. The  $\alpha$ -1,2-fucosyltransferase function of *wbsJ* gene was unambiguously determined by enzymatic assay and NMR spectroscopy analysis. We believe that the elucidation of O-antigen biosynthesis cluster sequence will help in the design of an O128 PCR serotyping test and the development of an O128 antigen-based, structurally defined glycoconjugate vaccine.

## References

- [1] Nataro, J.P. and Kaper, J.B. (1998) Clin. Microbiol. Rev. 11, 142–201.
- [2] Trabulsi, L.R., Keller, R. and Tardelli Gomes, T.A. (2002) Emerg. Infect. Dis. 8, 508–513.
- [3] Bortolini, M.R., Trabulsi, L.R., Keller, R., Frankel, G. and Sperandio, V. (1999) FEMS Microbiol. Lett. 179, 169–174.
- [4] Robins-Browne, R.M., Yam, W.C., O'Gorman, L.E. and Bettelheim, K.A. (1993) J. Med. Microbiol. 38, 222–226.
- [5] Nataro, J.P., Baldini, M.M., Kaper, J.B., Black, R.E., Bravo, N. and Levine, M.M. (1985) J. Infect. Dis. 152, 560–565.
- [6] Giammanco, A., Maggio, M., Giammanco, G., Morelli, R., Minelli, F., Scheutz, F. and Caprioli, A. (1996) J. Clin. Microbiol. 34, 689–694.
- [7] Lim, Y.S., Ngan, C.C. and Tay, L. (1992) J. Trop. Med. Hyg. 95, 339–342.
- [8] Raetz, C.R. and Whitfield, C. (2002) Annu. Rev. Biochem. 71, 635–700.
- [9] Reeves, P.R. et al. (1996) Trends Microbiol. 4, 495–503.
- [10] Jennings, H. (1992) J. Infect. Dis. 165 (Suppl. 1), S156–S159.
- [11] Benaissa-Trouw, B., Lefeber, D.J., Kamerling, J.P., Vliegenthart, J.F., Kraaijeveld, K. and Snippe, H. (2001) Infect. Immun. 69, 4698–4701.
- [12] Anderson, P.W. et al. (1989) J. Immunol. 142, 2464–2468.
- [13] Bolgiano, B., Mawas, F., Yost, S.E., Crane, D.T., Lemercinier, X. and Corbel, M.J. (2001) Vaccine 19, 3189–3200.

- [14] Wessels, M.R., Paoletti, L.C., Guttormsen, H.K., Michon, F., D'Ambra, A.J. and Kasper, D.L. (1998) *Infect. Immun.* 66, 2186–2192.
- [15] Sengupta, P., Bhattacharyya, T., Shashkov, A.S., Kochanowski, H. and Basu, S. (1995) *Carbohydr. Res.* 277, 283–290.
- [16] Sengupta, P., Bhattacharyya, T., Majumder, M. and Chatterjee, B.P. (2000) *FEMS Immunol. Med. Microbiol.* 28, 133–137.
- [17] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) *Nucleic Acids Res.* 22, 4673–4680.
- [18] Hofmann, K. and Stoffel, W. (1993) *Biol. Chem.* 374, 166.
- [19] Ginsburg, V. (1961) *J. Biol. Chem.* 236, 2389–2393.
- [20] Frick, D.N., Townsend, B.D. and Bessman, M.J. (1995) *J. Biol. Chem.* 270, 24086–24091.
- [21] Alexander, D.C. and Valvano, M.A. (1994) *J. Bacteriol.* 176, 7079–7084.
- [22] Wang, L., Huskic, S., Cisterne, A., Rothmund, D. and Reeves, P.R. (2002) *J. Bacteriol.* 184, 2620–2625.
- [23] Parolis, H., Parolis, L.A. and Olivieri, G. (1997) *Carbohydr. Res.* 303, 319–325.
- [24] Orskov, F., Orskov, I., Sutton, A., Schneerson, R., Lin, W., Egan, W., Hoff, G.E. and Robbins, J.B. (1979) *J. Exp. Med.* 149, 669–685.
- [25] Berry, D.S., Lynn, F., Lee, C.H., Frasch, C.E. and Bash, M.C. (2002) *Infect. Immun.* 70, 3707–3713.
- [26] Katzenellenbogen, E. et al. (1998) *Eur. J. Biochem.* 251, 980–985.
- [27] Yao, Z., Liu, H. and Valvano, M.A. (1992) *J. Bacteriol.* 174, 7500–7508.
- [28] Breton, C., Oriol, R. and Imberty, A. (1998) *Glycobiology* 8, 87–94.
- [29] Shao, J., Zhang, J., Kowal, P., Lu, Y. and Wang, P.G. (2002) *Biochem. Biophys. Res. Commun.* 295, 1–8.